《心理学报》论文自检报告

请作者填写以下内容，粘贴在稿件的首页。

1. 请以"研究亮点"的形式列出最多三条本研究的创新性贡献，总共不超过 200 字。
《心理学报》的目标是发表"既科学优秀，又具有广泛兴趣和意义" (be both scientifically excellent and of particularly broad interest and significance)的心理学前沿研究。如果您的研究只有小修小补的贡献，没有尝试开创新的研究领域(new areas of inquiry)或提出独到见解和创新视角(unique and innovative perspectives)，特别纯粹只是研究没有明确心理学问题的算法或技术的工作，这类研究被本刊接受的机会小，建议另投他刊。
答:

a. 从音乐到语言加工的方向，首次证明了音乐音高加工经验可以全面提高普通话母语音乐家的音系能力;
b. 延伸了音乐与语言加工的"共享论": 在超音段层面上证明了音乐与语言经验不仅可以在一般性声学层面发生迁移，语言领域的特异性加工经验——声调范畴感知也和音乐音高加工经验存在相互影响。

2. 作者已经投稿或发表的文章中是否采用了与本研究相同的数据？如果是，请把文章附上审查。(我们不赞成作者用同一数据发表多篇变量相同的文章，也不赞成将一系列的相关研究拆成多个研究来发表的做法。)
答: 否。

3. 管理、临床、人格和社会等领域仅有自我报告(问卷法)的非实验非干预研究，需要检查数据是否存在共同方法偏差(common method bias)。为控制或证明这种偏差不会影响研究结论的效度，你使用了什么方法？采取了哪些措施？(共同方法偏差的有关文献可参见: http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract894.shtml)基于横断数据，仅有自我报告，仅仅在方便样本中施测，这样的研究数据易取得，但通常创新性价值不大，被本刊接受的机会小。
答: 本研究为实验研究，无问卷设计。

4. 是否报告并分析了效果量(effect sizes; 如: *t* 检验: Cohen's *d*; 方差分析: η^2 或 η_p^2; 标准化回归系数)？(很多研究只是机械地报告了效果量，但没有做必要的分析或说明，如效果量是大中小？有什么理论意义或应用意义？)。(在 google 中搜索"effect size calculator"，可搜到许多计算方便的 APP。效应量的有关解释，中文可参考: http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract1150.shtml；英文可参看: http://www.uccs.edu/lbecker/effect-size.html
是否报告统计分析的 95% CI？(如，差异的 95% CI; 相关/回归系数的 95% CI)置信区间的有关计算和绘图可参考 https://thenewstatistics.com/itns/esci/)
答: 已报告效应量和统计分析的 95% CI。

5. 请写出计划的样本量，实际的样本量。如果二者有差别，请写出理由。以往心理学研究中普遍存在样本量不足导致的低统计功效(power)问题，我们建议在论文的方法部分解释您计算及认定样本量的依据。应该以有一定依据的效果量(effect size)、期望的功效来确定样本量，并报告计算用软件或程序。样本量计划的理由和做法可参考 https://osf.io/5awp4/
答: 根据本研究参数的设定（已在文章中汇报），需要样本量为 34，参考以往领域内文章，实际样本量设定为 60。

6. 假设检验中，如果是零假设显著性检验(NHST)，需报告精确 *p* 值而不是 *p* 的区间(小于 0.001 的报告区间，其他报告精确 *p* 值)。你的论文是否符合该项要求？如果是贝叶斯因素，是否已报告其对先验分布假定的敏感性？
答: 符合该项要求。

7. 为保证论文中数据报告的完备性，统计分析中如果剔除了部分数据，是否在文中报告？原因是什么？包含这部分数据的统计结果如何变化？统计分析中是如何处理缺失数据的？使用量表时是否删除了其中的个别题目？原因是什么？如果包含这部分题目，统计结果会如何变化？是否有测量的项目或者变量没有报告？原因是什么？请写出在论文中的位置。
答: 本研究已满足数据报告的完备性。

8. 研究用到的未经过同行评议和审查的实验材料、量表或问卷, 是否附在文件的末尾以供审查？如果没有, 请写出理由。如果该文发表, 您是否愿意公开这些材料与其他研究者共享？
答: 本研究不涉及未经过同行评议和审查的实验材料、量表或问卷。

9. 本刊要求作者提供原始数据, 请在以下 3 种里选择一种打√:

a)投稿后将数据发至编辑部邮箱 (√)
b)数据可以从如下链接中获得 ____ ( )
c)原始数据和程序已在科学数据银行(https://www.scidb.cn/surl/xlxb)上分享 ( )
d)如不能提供，请说明理由或提供有关证明。

10. 您的研究是否是临床干预或实验室实验？ 是( ) 否(√)
如果是, 请提供预注册登记号 ___________。
如果没有, 请说明原因____________________。
注: 临床干预或实验室实验, 建议在收集数据前预注册(pre-register)。也鼓励其他实验研究预注册。预注册要求写出所有的研究假设及其支持, 以及实验/干预的详细过程和步骤。本期刊的预注册网站是 https://os.psych.ac.cn/preregister (使用说明书见本刊网站"下载中心")或 https://osf.io/ 或 https://aspredicted.org/。如果您的研究有预注册, 会显著增加被录用的机会。预注册的重要性可参考 https://osf.io/5awp4/

11. 您的研究如果用到了人类或动物被试, 是否得到所在单位伦理委员会的批准？如果是, 请把扫描版发至编辑部邮箱。如果否, 请说明理由。
答: 是，已发送伦理审查文件。

12. 是否依据编辑部网站发布的"英文摘要写作注意事项"撰写 400~500 个单词的英文大摘要？英文题目和摘要是否已请英语好的专业人士把关或者已送专业 SCI/SSCI 论文编辑公司修改润色？
答: 是，已润色。

13. 如果第一作者是学生, 请导师单独给编辑部(xuebao@psych.ac.cn)发邮件, 说明已阅读本文并认真把关。是否已提醒导师给编辑部发邮件？(编辑部收到导师邮件后才会考虑进入稿件处理流程)
答: 第一作者是学生，已提醒导师发送邮件。

14. 请到编辑部网站首页右侧"下载中心"下载并填写"稿件不涉密证明", 加盖通讯作者单位的保密办公章, 把扫描件发至编辑部邮箱(xuebao@psych.ac.cn)。如没有保密办公章, 请加盖通讯作者的单位公章。是否已发邮件？
答: 是。

语言与音乐经验存在双向迁移：基于普通话母语音乐家声调范畴感知的研究

摘 要 研究构造普通话声调 T1-T2 连续统, 采用声调范畴感知识别与区分任务, 考察: (1)音乐经验能否影响普通话母语音乐家的音系能力; (2) 普通话母语音乐家的声调范畴感知模式能否迁移至精细化的音乐加工当中。结果发现: (1) 普通话母语音乐家具有更高的声调范畴感知程度, 在识别曲线陡峭度、范畴边界宽度、范畴间区分率、区分峰度等指标上均有体现; (2)普通话母语音乐家在音乐刺激的识别与区分曲线上均表现出范畴感知模式。研究表明, 音乐经验能够跨领域提升普通话母语音乐家的音系能力, 声调范畴感知模式能够跨领域迁移到音乐感知中。研究结果证明普通话母语音乐家语言与音乐经验存在双向迁移, 从声调范畴感知这一领域特异性角度支持了"训练迁移效应"。

关键词 语言, 音乐, 声调范畴感知, 训练迁移效应

分类号 B842

1 引言

语言(speech)和音乐(music)是人类两种用于交际的产物, 两者存在着紧密联系, 均涉及复杂的认知过程, 两个领域的加工均调用音高信息(南云, 2017; Nan & Friederici, 2013; Patel, 2008)。音高指在知觉范畴内声音的"高""低", 由基频(F0)决定。在音乐中, 音高独特的排列构成乐曲的旋律(Bakan, 2012)。而在语言中, 音高可以区分语义, 一般由单音节上的音高轮廓模式的变化来实现, 具有这种特征的语言被称作声调语言(Wang, 1976)。

1.1 "训练迁移效应"与声调范畴感知

基于语言和音乐加工资源调用的相似性, 研究者提出音乐与语言加工的"共享论": 认为语言和音乐加工具有共同的神经生理基础, 二者共享声音范畴学习的认知和神经生理机制(Patel, 2003, 2008; Friederici, 2017)。Besson 等(2011)进一步指出, 语言和音乐共享神经加工机制, 两个领域不仅仅会在领域一般性的普通声学层面发生迁移, 音乐和语言领域的经验也可以在领域特异性层面出现迁移, 体现了"训练迁移效应(transfer of training effects)"。

在领域一般性层面, 相关研究证实了语言和音乐中的音高加工经验存在双向迁移(Cooper & Wang, 2012; Giuliano et al., 2011; Marie et al., 2011; Zhao & Kuhl, 2015)。例如, Marie 等(2011)发现法语母语音乐家对汉语普通话(以下简称"普通话")声调辨别的精度显著高于非音乐家, 证明了音乐到语言方向的迁移; Giuliano 等(2011)发现普通话母语者(声调语言母语者)音乐音高的表征精度相对于非声调语言母语者明显提高, 证明了语言到音乐方向的迁移。然而, 除了领域一般性的声学资源, 音乐和语言加工还需要调动领域特异性资源。语言加工需要调动语言领域特异的音系(phonology)资源。音系指某一特定语言(方言)的语音系统, 其成分是该语言中能够区别意义的最小发音单位, 即音位(phoneme)。音系水平上的加工指的是对某种语言内音位的识辨, 音系能力越强, 越能有效区别语言中具有区别意义作用的最小语音单位(刘富华, 孙炜, 2009)。研究表明, 由于长期的声调语言经验, 声调语言母语者在长期音系记忆中建立起了可靠的音系表征模板, 对母语曲拱声调形成范畴感知模式(Francis et al., 2003; Liberman et al., 1957; Wang, 1976; Xu et al., 2006; 张林军, 2010)。与音乐音高的连续感知不同, 在语言的范畴感知模式下, 声调连续统中连续渐变的音高会被感知为离散的、有限的音系表征, 声调母语者需要忽略音系范畴内不同音高的声学差异, 而对跨范畴边界的声学差异保持敏感(Liberman et al., 1957), 这便是音系能力的体现。

1.2 音乐经验对声调范畴感知的影响

当前, 已有研究在音段层面上证实了音乐经验对音系能力的影响, 支持"训练迁移效应"。例如, Marie 等(2011)发现音乐家区分音节差异的能力显著强于非音乐家; Chobert 等(2011)发现音乐家儿童对语音中发声起始时间(voice onset time, VOT)的细微差异更为敏感; Bidelman 等(2014)发现音乐家在感知元音变化时更为准确且迅速。这些研究证实, 音乐家处理音段上的音系差异也具有一定优势, 说明低水平感觉层面上的能力提升也可能对高层次的认知构建(例如语言的音系表征)产生影响(Chen & Peng, 2020; Besson et al., 2011)。

但是, 在超音段层面, 研究还不充分。有关训练迁移效应的研究中, Tang 等(2016)发现了普通话母语的音乐家在感知语言声调变化时表现出更大的失匹配负波。然而, 音乐家对声调感知的增益可能来自于声学和音系两方面, 这一研究没有很好的剥离声学与音系加工。经典声调范畴感知实验采用识别和区分任务, 要求被试对来自不同范畴的刺激进行识别或区分, 通过计算可以分别对声学和音系加工水平进行测量, 从这两个维度探讨音乐经验向语言加工的迁移(Wu et al., 2015; Xu et al., 2006)。以往通过声调语言母语者考察成人音乐经验对母语

范畴感知影响的研究仅有三项(Chen et al., 2020; Wu et al, 2015; Zhu et al, 2021), 且均未能发现音乐家在音系能力上的优势, 只发现了音乐经验在一般声学能力上的迁移。已有研究对声调连续统的语言学属性认识不足, 未能提供合适的实验材料, 这可能是未发现音乐家音系能力优势的原因。Zhu 等(2021)采用的 T2-T4[1]连续统声调走向差异和刺激间隔都过大, 易导致对连续统的感知出现天花板效应。Wu 等(2015)、Zhu 等(2021)和 Chen 等(2020)采用的刺激时长过短, 不能诱发出正常语音感知下的范畴感知模式。

1.3 声调范畴感知模式对音乐感知的影响

借助范畴感知范式也可以观察语言经验向音乐感知的迁移(Chang et al, 2016; 李贤卓等, 2022)。感知声调语言时, 范畴感知模式需要母语者忽略范畴内的声学差异，将音高感知为离散的音系表征; 而感知音乐时, 则要求尽可能对音高信息进行精细化加工, 准确感知音符的音高, 进行连续化的感知。二者之间对于音高精细化加工程度的要求相反。结合“共享论”, 这提示声调范畴感知经验可能会对音乐音高精细化加工造成阻碍(李贤卓等, 2022; Patel, 2014)。目前, 已经有证据证明了这一假设。Chang 等(2016)发现, 普通话母语非音乐家在音乐旋律区分任务上的表现不如英语母语音乐家与非音乐家, 说明语言的声调范畴感知经验可能迁移到了音乐感知中, 削弱对音乐音高变化的敏感程度。李贤卓等(2022)发现声调母语者的范畴感知模式可以跨领域迁移到音乐上, 在音乐感知上也形成了类似的范畴感知模式, 阻碍了音乐的精细化加工。以上两项研究均考察的是普通话母语非音乐家, 而普通话母语音乐家在声调语言和音乐两个领域均有长期经验, 同时具备精细化的声学感知和音系层面的范畴感知模式。基于音乐和语言加工“共享论”的假设和上述证据, 声调母语音乐家两个领域的加工经验可能存在冲突, 或存在着更为复杂的加工迁移模式, 仍有待进一步研究。

1.4 研究问题与假设

综上所述, 本研究拟针对普通话母语音乐家人群, 考察: (1)音乐经验能否影响普通话母语音乐家的音系能力, 从而调节其声调范畴感知模式; (2)普通话母语音乐家的声调范畴感知模式如何迁移至精细化的音乐加工当中。

针对以上两个研究问题, 本研究拟使用普通话 T1-T2 连续统, 构造对应的语言和音乐刺激, 对比音乐家、非音乐家在语言、音乐刺激上的范畴感知情况。我们预期音乐家的范畴感

[1] T1、T2、T3、T4 分别代表普通话一声(阴平)、二声(阳平)、三声(上声)、四声(去声)。

知程度大于非音乐家, 可从音乐到语言的方向证明音乐经验向语言特异性加工领域的迁移; 同时, 音乐家在感知音乐刺激时会出现和语言类似的范畴感知模式, 从音乐到语言的方向验证语言领域特异性范畴感知向音乐加工的迁移效应。通过两个方向的考察, 本研究将全面研究普通话母语音乐家语言与音乐经验的双向迁移效应, 从语言特异性音系加工的角度验证“训练迁移效应”。

2 研究方法

2.1 被试

采用 G*Power 3.1.9.7(Faul et al., 2007)估算实验样本量, 设置参数 effect size f = 0.25、α err prob = 0.05、power(1 – β err prob) = 0.80, 计算出所需样本量为 34 人。参考领域内前人研究(李贤卓等, 2022; Wu et al., 2015; Zhu et al., 2021), 实验最终共招募 60 名被试(N = 60)。被试分为音乐家组(n = 30)与非音乐家组(n = 30)。其中音乐家组男 16 人, 女 14 人, 平均年龄 19.30 岁(SD = 1.49), 受专业音乐训练时长不少于 7 年, 且训练起始年龄不晚于 12 岁, 被试所修专业均为音乐相关(按照 Wong et al., 2007; Wayland et al., 2010 的标准)。非音乐家组男 16 人, 女 14 人, 平均年龄 22.10 岁(SD = 2.14), 所有非音乐家组的被试在近 5 年内没有受过任何音乐训练, 若 5 年前曾受过音乐训练, 总时长不多于 2 年(按照 Wong et al., 2007; Wayland et al., 2010 的标准)。采用蒙特利尔失歌症评估测试(Montreal battery of evaluation of amusia, MBEA; Peretz et al., 2003)评估被试是否有失歌症, 两组被试中均未发现失歌症者。所有被试均为右利手, 无听力、语言或神经功能等障碍, 智商正常, 母语均为普通话或北方方言[2]。所有被试均通过网络海报招募而来, 在参与实验前均签署知情协议, 参与实验后均获得一定的报酬。

2.2 材料

本研究的材料基于音节[i], 构造从 [i˥](T1, 普通话阴平调)到[i˧˥](T2, 普通话阳平调)等声学间距变化的 T1-T2 连续统。相比于其他连续统, T1-T2 连续统有下面几个优势。首先, T1-T2 连续统跨度较小(相较于 T2-T4 连续统), 对连续统中刺激的识别与区分有一定的难度, 该

[2] 为了最大限度避免方言的干扰, 被试优先选择家庭交流语言为普通话的普通话母语者。所有被试的普通话水平均在二级甲等及以上。

连续统有潜力观测出音乐家范畴感知能力的加强。其次, T1-T2 连续统中刺激的变化单纯是 F0 的变化, 对其识别与区分的好坏直接取决于音系能力和声学能力, 受到其他因素的干扰较少。而对其他连续统, 例如对 T2-T3 连续统中的刺激的识辨常与声调转折的时间点有关

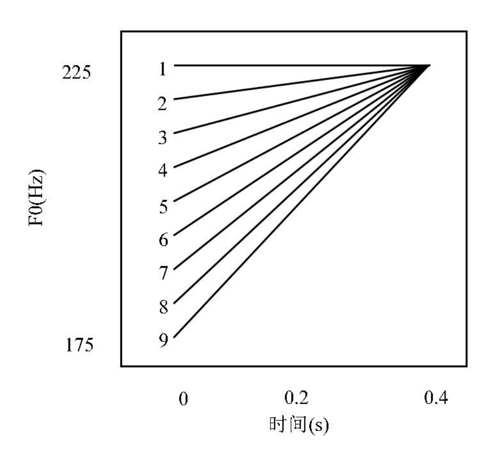


图 1 T1-T2 连续统音高模式示意图

(Shen & Lin, 1991), 且会受到普通话连续音变模式的干扰(Hao, 2012)。

实验材料含有语言刺激和音乐刺激两种刺激类型。制作连续统的原始音节为单音节[i[1]], 单音节样本由一位普通话二级甲等的男性朗读, 无上下文语境。使用 Praat 软件(Boersma & van Heuven, 2001), 首先将原始单音节样本起点和终点的 F0 均调整至 225 Hz, 作为刺激 1(T1)。然后保持终点的 F0 不变, 将起点 F0 降低至 175 Hz, 形成刺激 9(T2)。接下来, 将刺激 1 与刺激 9 作为边界, 在内部构造出等 F0 间隔[3]的其余 7 个刺激, 最终形成一个包含 9 个等间隔刺激的 T1-T2 连续统(如图 1 所示; 连续统详细音高参数见表 1)。单个刺激的时长为 400 ms, 与普通话阴平调、阳平调单音节的真实发音时长接近(郭锦桴, 1993)。再使用 Praat 软件中的音高提取功能, 分别提取连续统中 9 个语言刺激的音高轮廓, 将其合成为正弦连续音以模拟音乐刺激, 制作音乐刺激连续统(参考 Xu et al., 2006)。这些正弦波与对应语言刺激的音高轮廓、振幅和持续时间相同, 主要声学差异是频谱含量。以刺激 9 为例, 该刺激在语言中被感知为普通话阳平调, 在音乐中近似于一个从 F3(采用科学音高记号法, 下同)到 A3 的滑音。图 2 呈现了语言连续统和音乐连续统中刺激 9 的波形图和频谱图。

表 1 连续统各条刺激的起始音高参数

刺激	起点音高(Hz)	起点音高(ERB)	步幅(Hz)	步幅(ERB)
1	225	6.229	6.63	0.12
2	218.37	6.104	6.52	0.12

[3] 在计算 F0 的步幅时, 将 F0 的单位转换为等效矩形带宽标度(equivalent rectangular bandwidth, ERB), 该指标能真实地衡量心理感知到的频率变化情况(Greenwood, 1961)。

3	211.85	5.980	6.41	0.12
4	205.44	5.856	6.30	0.12
5	199.14	5.732	6.19	0.12
6	192.95	5.607	6.09	0.12
7	186.87	5.483	5.98	0.12
8	180.88	5.359	5.88	0.12
9	175	5.235		

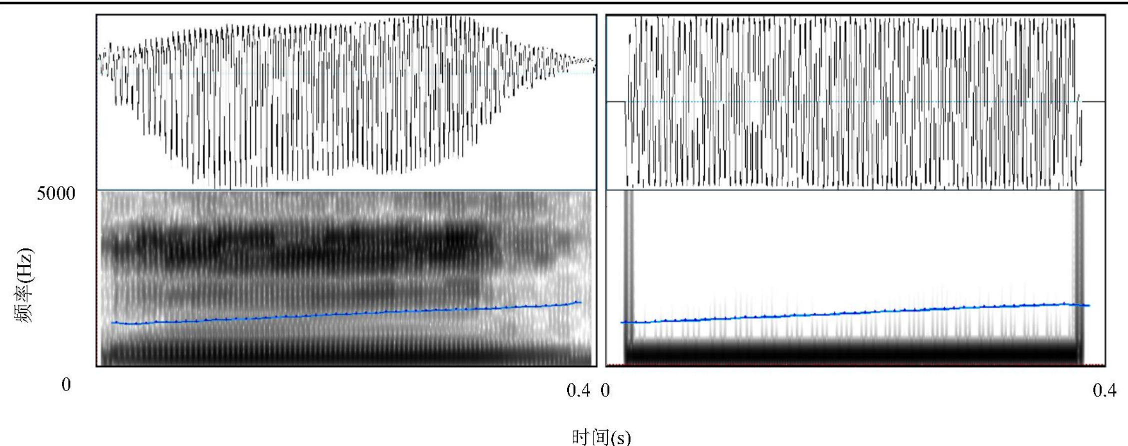


图 2 语言连续统和音乐连续统中刺激 9 的波形图和频谱图

（左上：语音刺激波形图；左下：语音刺激频谱图；右上：音乐刺激波形图；右下：音乐刺激频谱图）

2.3 过程

实验通过 E-prime 3.0 在计算机上向被试呈现。被试坐在一个安静的密闭环境中, 佩戴耳机按照要求进行按键反应, 耳机双声调播音, 两个声道的音量保持相等且一致, 音量设定为 80 dB。

识别任务中, 首先在屏幕中央出现十字注视点 500 ms, 随后播放一个语言或音乐刺激, 刺激播放完毕后呈现带有指示语的反应屏, 被试需判断该刺激为普通话 T1 或 T2 并按键。被试最多有 2500 ms 的时间进行判断, 按键后反应屏随即消失。反应屏消失后呈现 500 ms 空白, 随后进入下一循环组 (如图 3 所示)。音乐刺激与语言刺激分别在两个组块(block)中呈现。每个组块中, 连续统中的 9 个刺激将各重复 10 次, 每一组块各呈现 90 个刺激, 组块内刺激呈现顺序随机。语言刺激组块与音乐刺激组块的呈现顺序随机平衡, 两个组块中间有短暂

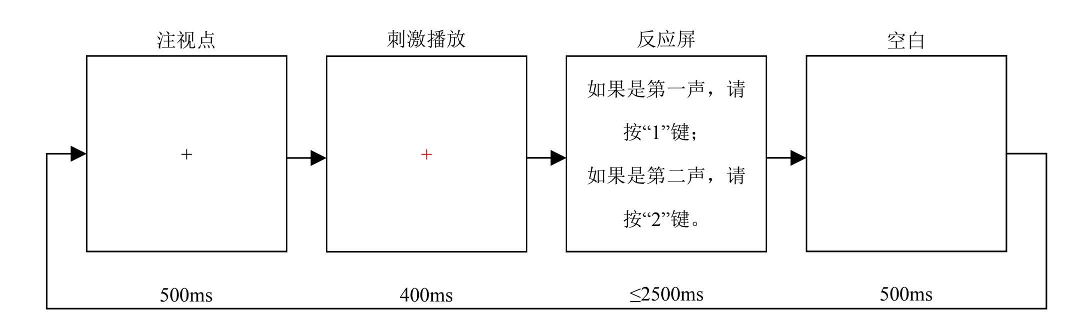


图 3 识别任务流程图

的休息时间。

区分任务中, 连续统中的两个刺激成对出现, 刺激间隔为 500 ms, 即一次完整的语言或音乐刺激对总时长为 1300 ms。实验开始后, 首先在屏幕中央出现十字注视点 500 ms, 随后播放一个语言或音乐刺激对, 刺激对播放完毕后呈现带有指示语的反应屏, 被试需判断刚播放的刺激对中的两个刺激是否相同并按键。被试最多有 2500 ms 的时间进行判断, 按键后反应屏随即消失。反应屏消失后呈现 500 ms 空白, 随后进入下一循环组(如图 4 所示)。音乐刺激与语言刺激分别在两个组块中呈现。每个组块中, “不同”刺激对(即刺激对中的两个刺激不同)中的两个刺激均为二阶差异, 存在正序与倒序两种组合顺序。例如, 刺激 1 与刺激 3 组合为一个刺激对, 先播放刺激 1, 再播放刺激 3, 形成刺激对 1-3; 将两个刺激的顺序颠倒, 形成刺激 3-1。按照相同的规则, 每一组块中均含有 1-3、3-1、2-4、4-2、3-5、5-3、4-6、6-4、5-7、7-5、6-8、8-6、7-9、9-7 十四种“不同”刺激对。另外, 每一组块还含有 1-1、2-2、3-3、4-4、5-5、6-6、7-7、8-8、9-9 九种“相同”刺激对。因此, 每个组块中“相同”与“不同”刺激对共计 23 种, 每种刺激对呈现十次, 共呈现 230 对刺激, 呈现顺序随机。语言刺激组块与音乐刺激组块的呈现顺序随机平衡, 两个组块中间有短暂的休息时间。

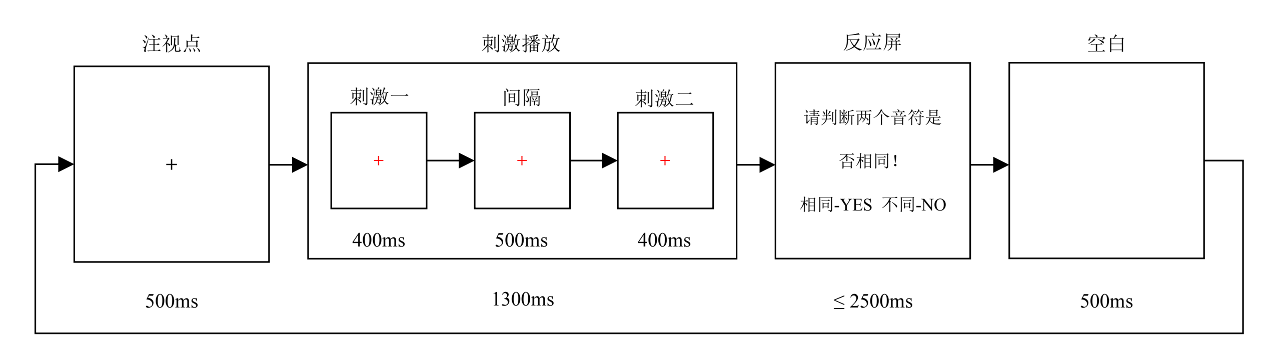


图 4 区分任务流程图

2.4 数据处理

实验计算与范畴感知相关的六个基本指标: 识别曲线陡峭度、范畴边界位置、范畴边界宽度、范畴间区分率、范畴内区分率、区分峰度(Jiang et al., 2012; Peng et al., 2010; Wu et al., 2015; Xu et al., 2006)。

识别任务结果采用 Xu 等(2006)提出的 logist 回归方程 $ln\frac{P_I}{1-P_I}=b_0+b_1$ 来对数据进行拟合, 以获得识别曲线(图 5)。P_I 代表被试将识别任务中的刺激判

断为 T1 的概率, x 代表刺激序号。我们用回归方程中系数b_1(即斜率)来作为指示识别曲线陡峭度的指标, b_1的绝对值越大说明识别曲线越陡峭, 范畴化程度越高。

另外, 我们令$P_I = 0.5$, 此时对应的x 值即为范畴边界位置, 得到$x_{cb} = -\frac{b_0}{b_1}$。依照 Peng 等(2010)的算法, 计算$P_I$ 分别取在 25%与 75%时, 横坐标x 距离的绝对值, 得到边界范畴宽度$W_{cb} = |\frac{2\ln 3}{b_1}|$, 较窄的范畴宽度是强范畴感知的体现。

然后, 根据区分任务中七个二阶刺激单元的区分正确率绘制区分曲线(图 7)。为了计算区分任务中的正确率, 我们将区分任务中每一组块的 230 个刺激对分为七个二阶刺激单元。每个 A-B 刺激单元(A 与 B 为跨度为二阶的刺激)包含四种刺激对(A-A、B-A、A-A、B-B)。例如, 1-3 刺激单元包含 1-3、3-1、1-1、3-3 四种刺激对。每个刺激单元共含有 40 个刺激对。相邻的两个刺激单元共享 AA 或 BB 刺激对(例如: 1-3 刺激单元与 3-5 刺激单元共享 3-3 刺激对)。每个刺激单元的区分正确率P_{AB} 按照如下的公式计算(Xu et al., 2006): $P_{AB} = P("S"|S) \bullet P(S) + P("D"|D) \bullet P(D)$。正确判断“相同”刺激对的概率(正确将 AA 与 BB 刺激对判断为“相同”的数量比上 AA 与 BB 刺激对的总数)与正确判断“不同”刺激概率(正确将 AB 与 BA 刺激对判断为“不同”的数量比上 AB 与 BA 刺激对的总数)分别用$P("S"|S)$与$P("D"|D)$表示。$P(S)$与$P(D)$为别为“相同”刺激对(AA 与 BB 刺激对)与“不同”刺激对(AB 与 BA 刺激对)占每个刺激单元刺激对总数的概率。按此方法计算, 每位被试在语言与音乐组块中可各得出P_{13}、P_{24}、P_{35}、P_{46}、P_{57}、P_{68}、P_{79}七个数据。

结合前文计算得出的范畴边界位置x_{cb} 的值, 我们取跨范畴边界位置刺激单元的正确率平均值记为范畴间区分率, 取其余单元的正确率平均值记为范畴内区分率(Jiang et al., 2012)。例如, 范畴边界$x_{cb} = 5.5$, 则取 4-6 刺激单元与 5-7 刺激单元区分正确率的平均值记为范畴间区分率P_{bc} , 除了这两个区分单元之外其他单元的区分正确率平均值记为范畴内区分率P_{wc} 。考虑到范畴内区分率反映的是声学层面上的区分能力, 而范畴间区分率与音系能力和声学能力均有关系, 为了进一步对指标进行纯化, 以便剥离出反

应音系能力的指标, 我们用范畴间区分率减去范畴内区分率, 得到区分峰度 $P_{pk} = P_{bc} - P_{wc}$, 区分峰度反映的是音系能力(Wu et al., 2015; Xu et al., 2006)。

3 结果

3.1 识别任务

图 5 呈现了各组的识别曲线。表 2 呈现了识别任务中的各指标值的描述统计结果。图 6 呈现了识别任务各指标的柱形统计图。

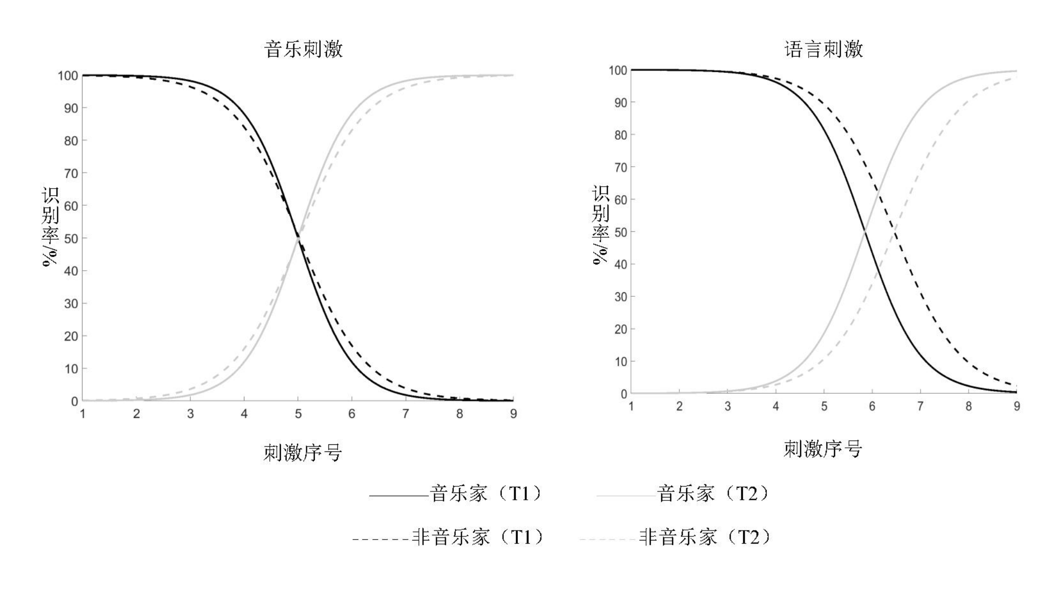


图 5 音乐家与非音乐家识别曲线

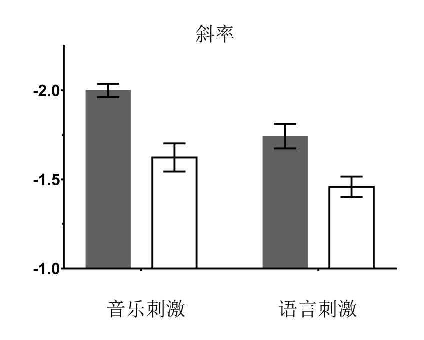
表 2 音乐家与非音乐家识别任务指标

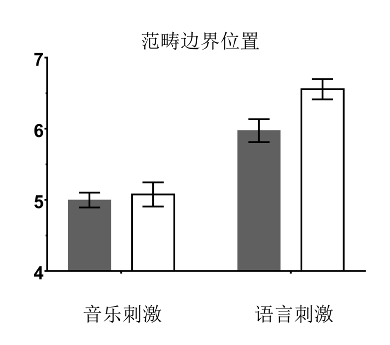
组别	音乐刺激			语言刺激		
	斜率	范畴边界位置	范畴边界宽度	斜率	范畴边界位置	范畴边界宽度
音乐家	-2.00(0.21)	5.00(0.57)	1.11(0.12)	-1.74(0.38)	5.97(0.88)	1.35(0.46)
非音乐家	-1.62(0.43)	5.08(0.93)	1.58(1.09)	-1.46(0.32)	6.56(0.78)	1.59(0.40)

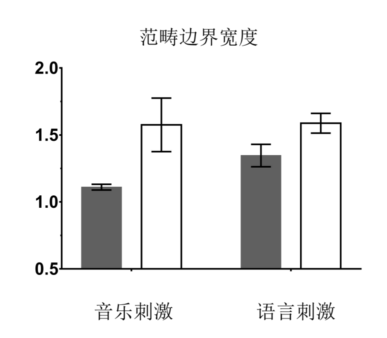
注: 括号内为标准差, 下同。

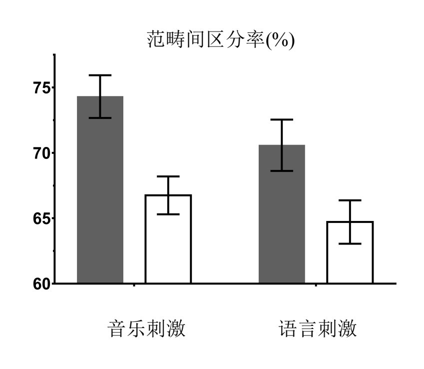
语言和音乐刺激的识别曲线(图 5)均显示, 跨范畴边界的两个相邻刺激的识别率差异远大于边界两侧的两个相邻刺激的识别率差异, 说明两组被试在语言和音乐刺激上都表现出了典型的范畴感知模式(Liberman et al., 1957; Xu et al., 2006)。对识别任务各指标(描述统计见

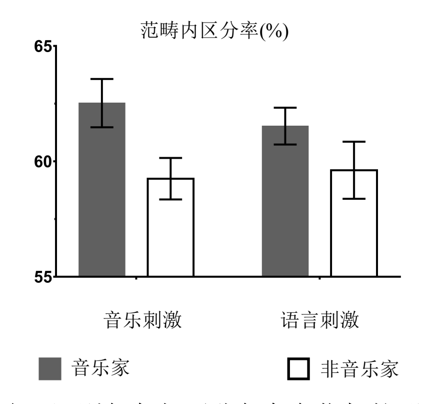
表 2)进行 2(组别: 音乐家, 非音乐家)×2(刺激类型: 音乐刺激, 语言刺激)重复测量方差分析。结果表明, 在识别曲线陡峭度上, 组别主效应显著, $F(1, 58) = 22.51$, $p < 0.001$, $\eta_p^2 = 0.28$, 95% CI = [-0.47, -0.19], 表现为音乐家识别曲线斜率显著小于非音乐家, 即音乐家识别曲线陡峭度显著高于非音乐家; 刺激类型主效应显著, $F(1, 58) = 14.69$, $p < 0.001$, $\eta_p^2 = 0.20$, 95% CI = [-0.32, -0.10], 表现为音乐刺激识别曲线斜率显著小于语言刺激, 即音乐刺激识别曲线陡峭度显著高于语言刺激。在范畴边界位置上, 组别主效应显著, $F(1, 58) = 5.36$, $p = 0.024$, $\eta_p^2 = 0.09$, 95% CI = [-0.62, -0.05], 表现为音乐家范畴边界比非音乐家显著靠近阴平端; 刺激类型主效应显著, $F(1, 58) = 66.69$, $p < 0.001$, $\eta_p^2 = 0.54$, 95% CI = [-1.53, -0.93], 表现为音乐刺激的范畴边界比语言刺激显著靠近阴平端。在范畴边界宽度上, 组别主效应显著, $F(1, 58) = 7.75$, $p = 0.007$, $\eta_p^2 = 0.12$, 95% CI = [-0.61, -0.10], 表现为音乐家的范畴边界显著窄于非音乐家组; 刺激类型主效应不显著, $F(1, 58) = 1.49$, $p = 0.227$, $\eta_p^2 = 0.03$, 95% CI = [-0.33, -0.08](图 6)。在所有指标上均未发现任何的交互效应(ps > 0.05)。











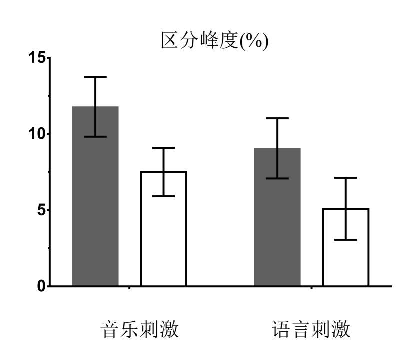


图 6 识别任务与区分任务各指标柱形统计图
(误差线表示标准误)

3.2 区分任务

图 7 呈现了各分组的区分曲线。表 3 呈现了区分任务中的各指标值的描述统计结果。图 6 呈现了区分任务各指标的柱形统计图。

表 3 音乐家与非音乐家区分任务指标

组别	音乐刺激			语言刺激		
	范畴间区分率	范畴内区分率	区分峰度	范畴间区分率	范畴内区分率	区分峰度
音乐家	0.74(0.09)	0.63(0.06)	0.12(0.11)	0.71(0.11)	0.62(0.04)	0.09(0.11)
非音乐家	0.67(0.08)	0.59(0.05)	0.08(0.09)	0.65(0.09)	0.60(0.07)	0.05(0.11)

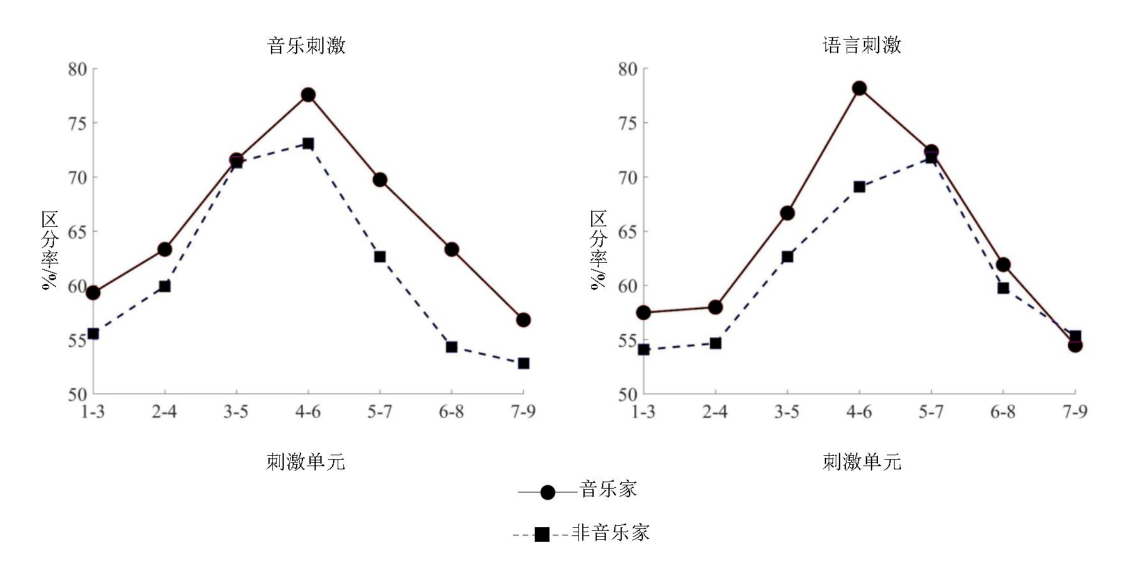


图 7 音乐家与非音乐家区分曲线

区分曲线(图 7)同样表明两组被试对音乐和语言的感知均为范畴感知, 具体体现为跨范畴边界刺激单元的区分率高于范畴同侧刺激单元的区分率(Liberman et al., 1957; Xu et al., 2006)。对区分任务各指标(描述统计见表 3)进行 2(组别: 音乐家, 非音乐家)×2(刺激类型: 音乐刺激, 语言刺激)重复测量方差分析。结果表明, 在范畴间区分率上, 组别主效应显著, $F(1, 58) = 14.02$, $p < 0.001$, $\eta_p^2 = 0.20$, 95% CI = [0.03, 0.10], 表现为音乐家范畴间区分率显著高于非音乐家。在范畴内区分率上, 组别主效应显著, $F(1, 58) = 4.95$, $p = 0.030$, $\eta_p^2 = 0.08$, 95% CI = [0.00, 0.05], 表现为音乐家范畴内区分率显著高于非音乐家。在区分峰度上, 组别主效应显著, $F(1, 58) = 4.95$, $p = 0.041$, $\eta_p^2 = 0.07$, 95% CI = [0.00, 0.08], 表现为音乐家区分峰度显著大于非音乐家(图 6)。所有刺激类型效应均不显著, 没有任何显著的交互效应($ps > 0.05$)。

4 讨论

本研究基于语言特异性声调范畴感知任务, 考察了普通话母语音乐家语言与音乐经验的双向迁移效应。识别任务中, 音乐家的识别曲线陡峭度显著大于非音乐家, 范畴边界位置较非音乐家显著靠近阴平端, 范畴边界宽度显著小于非音乐家; 在区分任务中, 音乐家的范畴间区分率、区分峰度均显著大于非音乐家。而且，音乐家与非音乐家在语言、音乐刺激的识别与区分曲线上均呈现明显的范畴感知模式。结果表明, 长期的音乐经验使普通话母语音乐

家的音系能力显著增强, 其语言范畴感知模式也跨领域迁移到了音乐感知中。尽管声调范畴感知具有领域特异性, 本研究仍借此观察到普通话母语音乐家感知语言和音乐时存在跨领域双向迁移效应。

4.1 音乐经验对普通话母语音乐家音系能力的提升

在识别曲线上, 更陡峭的范畴边界和更窄的范畴边界宽度指示范畴感知程度的提升(Jiang et al., 2012; 李贤卓等, 2022; Peng et al., 2010; Xu et al., 2006; 张林军, 2010)。本研究的结果表明音乐家较非音乐家的声调范畴感知程度更高, 音乐家的音系能力更强。在区分曲线上, 一般认为范畴内区分率体现的是声学信息加工的能力, 而范畴间区分率同时由声学加工能力与音系加工能力来决定(Wu et al., 2015; Xi et al., 2010; Yu et al., 2014; Zhang et al., 2011; Zhu et al., 2021)。用范畴间区分率减去范畴内区分率得到区分峰度, 进一步提取音系能力指标(Wu et al., 2015; Xu et al., 2006)。结果发现, 音乐家的范畴间区分率、范畴内区分率、区分峰度均显著高于非音乐家, 这指示音乐家的音系能力和声学能力均强于非音乐家。

在音乐经验向语言加工的迁移方向上, 前人多关注音乐和语言加工经验在一般性声学加工中的迁移效应, 证明了语言和音乐音高加工共享一般性声学加工资源(Cooper & Wang, 2012; Marie et al., 2011; Marques et al., 2007)。本研究从语言领域特异的音系加工层面入手, 首次发现音乐经验能够跨领域提升超音段音系的加工能力。本研究的这一发现验证了 Besson 等(2011)提出的训练迁移效应, 即音乐和语言领域的经验可以在领域特异性层面发生迁移。具体而言, 音乐经验带来的领域一般性能力的提升也可以对语言领域特异性能力产生影响。姚尧和陈晓湘(2020)的纵向研究曾发现, 学龄前幼儿在经过为期一年的音乐训练后, 其范畴边界宽度显著缩短, 从而证实音乐训练可以提高幼儿声调范畴感知程度。本研究的结果证明这一增益在成人上同样存在, 并通过进一步优化的材料在更多指标上发现了更为全面的促进效果。对此, 我们给予以下两种可行的解释: 首先, 在加工脑区上, 初级听觉皮层以及其附近的颞平面和极平面等区域负责言语感知早期的音位识别(Diesch et al., 1996; Levy & Wilson, 2020; Poeppel et al., 1997)。而这样的加工系统并非语言专职, 而必须处理各种类型的听觉输入(Friederici, 2017)。我们推测, 音乐训练可以增强负责分离和匹配时频模式的颞平面区域的加工能力, 从而提升其对声音信号时频特征分析的效率(Griffiths & Warren, 2002), 并将此增益迁移至语言音位感知, 带来音乐家音系能力的提升。Bidelman 等(2014)曾发现音乐家的初级听觉加工结构能更有效地提取与言语信息有关的谱频线索(如 F0、F1 等), 最终

使声音信号的表征更为清晰，这一研究为我们的推测提供了佐证。其次，音乐家范畴感知能力的提升可能与音义映射能力的增强有关。Bidelman 等(2014)指出音乐家在听觉通路中对语音信号具有的更稳健、更有选择性的内化表征；经过强化的表征为控制语音识别的决策过程提供了更可靠的音位模板，从而使经过提取的声学信号能够更准确、更迅速地通达存储在词典(lexicon)中的音系信息(Friederici, 2017)。

然而，本研究的这一发现与 Wu 等(2015)、Zhu 等(2021)、Chen 等(2020)的结论均相左，他们均未发现音乐经验对音系能力的影响。我们推测，这可能和连续统的选择有关，声调连续统的设计会影响声调的范畴感知模式：连续统的设计不同，范畴感知的结果会产生差异(吴倩, 王韫佳, 2018)。以上三项研究选择的连续统主要存在的问题有：第一，采用的连续统类型使跨范畴感知过易。Zhu 等(2021)在实验材料上选择了 T2-T4 连续统，T2-T4 连续统实际上跨越了两个范畴边界，实为 T2-T1-T4 连续统，内部声调走向差异过大，涉及曲拱调和平调的感知差异，导致对连续统的识别与区分很可能出现天花板效应。第二，连续统 F0 跨度过大。Zhu 等(2021)采用的连续统跨越了 100 Hz 以上，使得对连续统的区分与识别过于粗糙，过大跨度的连续统很有可能掩盖音乐家音系感知能力的优势，同样出现天花板效应。第三，连续统刺激时长过短。根据郭锦桴(1993)，保持阴平调、阳平调自然度的最低必要时长分别为 210 ms 与 230 ms，低于此界限的时长会使声调失去其音位特征。Wu 等(2015)、Zhu 等(2021)和 Chen 等(2020)采用的刺激时长均为 200 ms，过短的感知时间可能更多地调用到了声学的感知资源，不能诱发出正常语音感知下的范畴感知模式。而我们采用的 T1-T2 连续统音高步幅较小，识别和区分难度较大，需要更为精细的加工，且采用的 400 ms 的刺激时长接近普通话中的单音调时长，所以音乐家在音系加工能力上的优势才得以显著出来。不同的连续统选择及参数设置对实验结果会有怎样的影响，是未来的研究需要进一步讨论的问题。

4.2 普通话母语音乐家声调范畴感知模式向音乐感知的迁移

从识别曲线和区分曲线上可以看到：无论是音乐家还是非音乐家，语言刺激和音乐刺激的识别曲线均在范畴边界附近呈现较高的陡峭度，区分曲线上均存在明显的区分峰。这一结果表明两组被试在和语言刺激保持相同 F0 范围的音乐刺激中，也呈现出了范畴感知模式。值得注意的是，范畴感知属于语言领域的特异性能力，这说明存在于长期记忆中的母语音系感知模式可以跨领域迁移至音乐感知中。

根据 OPERA-e 假说(Patel, 2014), 音乐比语言在领域共享的声学加工中需要更高的精度。对音高加工而言, 音乐的连续性感知需要更高的加工精度, 以识别连续变化的音高线索, 属于细粒化(fine-grained)加工; 而语言中, 声调在音高上变化不及音乐那样精细, 对声调的感知无需对范畴内的音高差异进行区分, 属于粗粒化(coarse-grained)加工。前人通过对声调母语非音乐家进行考察, 发现粗粒化的语言音高范畴加工会向细粒化的音乐音高加工迁移, 从而会阻碍音乐音高的精细化加工(Chang et al., 2016; 李贤卓等, 2022)。普通话母语音乐家在具有声调范畴长期记忆的同时, 也具有更为精细的声学加工。这说明在语言范畴感知可能阻碍音乐音高精细化加工的同时, 他们的音乐经验又要求他们保持对音乐音高的细粒化加工。本研究的结果表明, 当语言范畴感知迁移至音乐音高加工中时, 音乐家对于音高的细粒化加工并未削弱其范畴感知程度, 表明普通话母语音乐家存在语言经验向音乐感知的迁移, 语言领域特异性范畴加工经验可以对一般声学加工产生影响。以往研究中, Wu 等(2015)同样发现音乐家在加工非语言刺激时激活了类似语言的范畴感知模式。这些结果证明, 语言音高范畴加工经验对声学加工的影响是牢固的, 这种语言领域的特异性感知模式可向其他领域的听觉加工迁移。相关脑科学研究发现, 处理音系信息的区域位于听觉加工系统更深层的结构中, 例如颞上沟和颞中回(Zhang et al., 2011)。据此, 我们推论, 语言范畴感知模式向音乐加工的迁移可能是由于母语音系加工经验的获得始于出生阶段, 音系记忆长期且稳定地存在于深层的认知资源中, 并对低层级的听觉输入产生影响(Wu et al., 2015; Xu et al., 2006)。

结合 4.1 的讨论, 本研究的结果表明, 普通话母语音乐家音高精细化加工和范畴感知模式同时存在, 且发生了跨领域(语言、音乐)共享。在范畴感知这一语言领域特异性加工的层面上, 普通话母语音乐家语言经验与音乐经验的相互迁移同时存在, 证实了语言和音乐加工经验带来的大脑可塑性变化是双向的(Bidelman et al., 2013)。

5 结论

本研究借助普通话声调范畴感知范式, 通过观察声调范畴感知表现, 考察普通话母语音乐家音乐和语言经验的双向迁移效应。研究发现, 音乐经验可以迁移至语言特异性感知当中, 提高了音系加工能力; 语言音系经验也可以迁移至音乐感知当中, 使普通话母语音乐家在音乐感知中呈现范畴化模式。研究结果证明普通话母语音乐家语言与音乐经验存在双向迁移

效应, 从声调范畴感知这一领域特异性角度支持了“训练迁移效应”(Besson et al., 2011)。研究支持语言与音乐加工的“共享论”(Giuliano et al., 2011; Patel, 2003, 2008; Xu et al., 2006)。

参考文献

Bakan, M. B. (2012). *World music: traditions and transformations (2nd ed.)*. New York: McGraw-Hill.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Frontiers in Psychology*, *2*, 94.

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for bidirectionality between the domains of language and music. *PLoS ONE*, *8*(4), e60676.

Bidelman, G. M., Weiss, M. W., Moreno, S., & Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *The European Journal of Neuroscience*, *40*(4), 2662–2673.

Boersma, P., & van Heuven, V. (2001). Speak and unSpeak with Praat. *Glot International, 5*(9–10), 341–347.

Chang, D., Hedberg, N., & Wang, Y. (2016). Effects of musical and linguistic experience on categorization of lexical and melodic tones. *Journal of the Acoustical Society of America*, *139*(5), 2432–2447.

Chen, F., & Peng, G. (2020). Reduced Sensitivity to Between-Category Information but Preserved Categorical Perception of Lexical Tones in Tone Language Speakers With Congenital Amusia. *Frontiers in psychology*, *11*, 581410.

Chen, S., Zhu, Y., Wayland, R., & Yang, Y. (2020). How musical experience affects tone perception efficiency by musicians of tonal and non-tonal speakers?. *PloS one*, *15*(5), e0232514.

Chobert, J., Marie, C., François, C., Schön, D., & Besson, M. (2011). Enhanced passive and active processing of syllables in musician children. *Journal of cognitive neuroscience, 23*(12), 3874–3887.

Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *The Journal of the Acoustical Society of America*, *131*(6), 4756–4769.

Diesch, E., Eulitz, C., Hampson, S., & Ross, B. (1996). The neurotopography of vowels as mirrored by evoked magnetic field measurements. *Brain and Language*, *53*(2), 143–168.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175–191.

Francis, A. L., Ciocca, V., & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Perception and Psychophysics*, *65*(7), 1029–1044.

Friederici, A. D. (2017). *Language in our brain: the origins of a uniquely human capacity*. Cambridge, MA: The MIT Press.

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in psychology*, *2*, 146.

Greenwood, D. D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America, 33*(10), 1344–1356.

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in neurosciences*, *25*(7), 348–353.

Guo, J. (1993). *An Explication and Investigation of Chinese Tones and Intonation.* Beijing: Beijing Language Institute Press.

[郭锦桴. (1993). *汉语声调语调阐要与探索*. 北京: 北京语言学院出版社.]

Hao, Y. C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, *40*(2), 269–279.

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2012). Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics. *Memory and Cognition*, *40*(7), 1109–1121.

Levy, D. F., & Wilson, S. M. (2020). Categorical encoding of vowels in primary auditory cortex. *Cerebral Cortex*, *30*(2), 618–627.

Li, X., Xiao, R., & Liang, D. (2022). The cross-domain influence of tonal categorical perception and tonal complexity on musical pitch perception. *Acta Psychologica Sinica, 54*, 1021–1030.

[李贤卓, 肖容, 梁丹丹. (2022). 声调范畴感知和声调复杂度对音乐音高感知的跨领域影响. *心理学报, 54*, 1021–1030.]

Liu, F., & Sun, W. (2009). *General theory of linguistics.* Beijing: Beijing Language and Culture University Press.

[刘富华, 孙炜. (2009). *语言学通论*. 北京: 北京语言大学出版社.]

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology*, *54*(5), 358–368.

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in mandarin Chinese. *Journal of Cognitive Neuroscience*, *23*(10), 2701-2715.

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *19*(9), 1453–1463.

Nan, Y. (2017). The facilitation effect of music learning on speech processing. *Advances in Psychological Science, 25*(11), 1844-1853.

[南云. (2017). 音乐学习对语言加工的促进作用. *心理科学进展*, *25*(11), 1844-1853.]

Nan, Y., & Friederici, A. D. (2013). Differential roles of right temporal cortex and Broca's area in pitch processing: evidence from music and Mandarin. *Hum Brain Mapp, 34*(9), 2045-2054.

Patel A. D. (2003). Language, music, syntax and the brain. *Nature neuroscience*, *6*(7), 674–681.

Patel, A. D. (2008). *Music, language, and the brain*. New York: Oxford University Press.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded opera hypothesis. *Hearing Research*, *308*, 98–108.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders——The Montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences, 999*, 58−75.

Peng, G., Zheng, H. Y., Gong, T., Yang, R. X., Kong, J. P., & Wang, W. S. Y. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics*, *38*(4), 616–624.

Poeppel, D., Phillips, C., Yellin, E., Rowley, H. A., Roberts, T. P. L., & Marantz, A. (1997). Processing of vowels in supratemporal auditory cortex. *Neuroscience Letters*, *221*, 145–148.

Shen, X. S., & Lin, M. (1991). A Perceptual study of Mandarin tones 2 and 3. *Language and Speech*, *34*(2), 145–156.

Tang, W., Xiong, W., Zhang, Y. X., Dong, Q., & Nan, Y. (2016). Musical experience facilitates lexical tone processing among mandarin speakers: behavioral and neural evidence. *Neuropsychologia*, *91*, 247–253.

Wang, W. S. (1976). Language change. *Annals of the New York Academy of Sciences, 280*(1), 61–72.

Wayland, R., Herrera, E., & Kaan, E. (2010). Effects of musical experience and training on pitch contour perception. *Journal of Phonetics*, *38*(4), 654–662.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422.

Wu, H., Ma, X., Zhang, L., Liu, Y., Zhang, Y., & Shu, H. (2015). Musical experience modulates categorical perception of lexical tones in native Chinese speakers. *Frontiers in Psychology*, 6, 436.

Wu, Q., & Wang, Y. (2018). Categorical perception of lexical tone and the neural mechanisms. *Advances in Psychological Science, 26*(01), 62–71.

[吴倩, 王韫佳. (2018). 声调的范畴知觉及其神经机制. *心理科学进展*, *26*(01), 62–71.]

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, *170*(1), 223–231.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America*, *120*(2), 1063–1074.

Yao, Y., & Chen, X. (2020). The effects of music training on categorical perception of Mandarin tones in 4- to 5-year-old children. *Acta Psychologica Sinica, 52*(04), 456–468.

[姚尧, & 陈晓湘. (2020). 音乐训练对 4～5 岁幼儿普通话声调范畴感知能力的影响. *心理学报, 52*(04), 456–468.]

Yu, K., Wang, R., Li, L., & Li, P. (2014). Processing of acoustic and phonological information of lexical tones in Mandarin Chinese revealed by mismatch negativity. *Frontiers in Human Neuroscience*, *8*, 729.

Zhang, L. (2010). On the influence of native language on the categorical perception of Chinese tones. *TCSOL Studies, 38*(2), 15–20.

[张林军. (2010). 母语经验对留学生汉语声调范畴化知觉的影响. *华文教学与研究*, *38*(02), 15–20.]

Zhang, L., Xi, J., Xu, G., Shu, H., Wang, X., & Li, P. (2011). Cortical dynamics of acoustic and phonological processing in speech perception. *PLoS ONE*, *6*(6), e20963.

Zhao, T. C., & Kuhl, P. K. (2015). Effect of musical experience on learning lexical tone categories. *Journal of the Acoustical Society of America*, *137*(3), 1452–1463.

Zhu, J., Chen, X., & Yang, Y. (2021). Effects of amateur musical experience on categorical perception of lexical tones by native Chinese adults: an ERP study. *Frontiers in Psychology*, *12*: 611189.

The bi-directional transfer between language and music experience: A study based on the tonal categorial perception of native Mandarin-speaking musicians

Abstract

The relations of music and language pitch processing have been extensively investigated during these years. Based on the similarity of resource invocation of language and music processing, researchers believe that language and music processing have a common neurophysiological basis and the experience of the two fields can be transferred to each other. The transfer effect of the domain-general acoustical experience of processing has been validated by a large number of studies. However, in addition to domain-general acoustic resources, language processing also involves phonological resources specific to the language domain, such as the categorical perception of tones. Researchers pointed out that experience in language and music can also transfer at the level of domain-specific competence, revealing the transfer of training effect. To date, whether this kind of transfer effect in terms of the categorical perception of tones happens remains unclear. Thus, the current study aimed to investigate the transfer of language and music experience in native Mandarin-speaking musicians in a bi-directional fashion.

The present study used a 2 (group: musicians vs. non-musicians) × 2 (stimulus type: music vs. speech) between-and-within-subjects design. Sixty participants were involved in the current study, including thirty native Mandarin-speaking musicians and thirty native Mandarin-speaking non-musicians. The identification and discrimination tasks in the traditional categorical perception paradigm were adopted, in which the Mandarin T1-T2 tonal continuum from [i˥] to [i˧˥] and its musical counterpart were constructed. The identification task required the participants to judge whether the stimulus in the continuum was T1 or T2, and the discrimination task asked participants to decide whether the two stimuli crossing two steps were the same or different.

Seen from the identification curve, musicians showed sharper and narrower categorical boundary compared to non-musicians. Both curves of musical stimuli and language stimuli yielded the typical pattern of categorical perception, i.e., the difference in identification rate

between the two adjacent stimuli across the boundary was much larger than that between the two adjacent stimuli on both sides of the boundary. In the discrimination task, musicians showed enhanced within-category discrimination accuracy, between-category discrimination accuracy, and discrimination peakedness. Also, the discrimination accuracy of between-category stimuli units could be seen higher than that of within-category stimuli units, which could be interpreted as a typical pattern of categorical perception.

The results showed that music experience could significantly enhance the phonological ability of native Chinese musicians, and their linguistic categorical perception pattern was transferred to musical perception. The conclusion could be drawn that there was a bi-directional transfer effect between the language and music experience of native Mandarin-speaking musicians on the categorical perception of tones, which provides empirical support for the "training transfer effect". The "shared domain-general view" of language and music processing, i.e., language and music processing share a common neurophysiological basis, could also be validated in terms of the phonological processing ability specific to the language domain.

Key words music; language; tonal categorical perception; transfer of training effect